Docket No. P18945

UTILITY PATENT

UNITED STATES APPLICATION FOR LETTERS PATENT

for

INTERCHANGEALBE CONNECTION ARRAYS FOR DOUBLE-SIDED DIMM PLACEMENT

by

Michael W. Leddige, Kuljit S. Bains, John T. Sprietsma

filed

March 2, 2004

INTERCHANGEABLE CONNECTION ARRAYS FOR DOUBLE-SIDED DIMM PLACEMENT

BACKGROUND

Currently, memory packages, such as dual, in-line memory modules (DIMM) may reside on both sides of a printed circuit board (PCB) or other substrate. This increases memory density for the system. Signals for the memories may route through the substrate, which may have several layers. With memory packages on both sides of the substrate, signal routing and integrity may become an issue.

Signals being routed through the substrate may connect to balls or connections to the DIMM on opposite sides of the packages. For examples, signals going to the closest side of the package on the one side of the substrate will generally end up having to go to the farthest side of the package on the other side of the substrate. The DIMM for the other side of the package is turned 'upside down' to mount it on the other side, causing the relevant connection to be on the other side of the package from the signal origination.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention may be best understood by reading the disclosure with reference to the drawings, wherein:

Figure 1 shows a prior art embodiment of a double-sided, dual, in-line memory module mounting.

Figure 2 shows an embodiment of a double-sided, dual, in-line memory module mounting.

Figures 3a and 3b show alternative arrangements of memory packages on a substrate.

Figure 4 shows an embodiment of a memory system using double-sided memory modules.

Figures 5a and 5b show alternative embodiments of stacked memory modules.

Figure 6 shows a connection diagram for an embodiment of a dual, in-line memory module having interchangeability.

Figure 7 shows a connection diagram for an alternative embodiment of a dual, in-line memory module.

Figure 8 shows a connection diagram for another alternative embodiment of a dual, in-line memory module.

Figure 9 shows a block diagram of an embodiment of a method to design dual, in-line memory module.

DETAILED DESCRIPTION OF THE EMBODIMENTS

Figure 1 shows a prior art embodiment of a double-sided, dual, in-line memory module (DIMM) mounting on a substrate. Double-sided refers to the fact that the memory modules are mounted on opposite sides of the substrate. The substrate 10 may be a multi-layered PCB, or any other substrate upon which DIMMs are mounted. Memory modules 12a and 12b are mounted on opposite sides of the substrate from each other, as are memory modules 14a and 14b. The substrate 10 has internally routed signal traces 16a and 16b. Stub 18 through via 13 allows the solder ball or other connection from the module 12a to connect to the first signal trace 16a.

The via 13 is manufactured on a larger pitch than the signal traces to afford extra width, and the use of several vias can limit the number of traces that can be routed through a single layer of the substrate 10. This may force additional layers and extra costs. In addition, to avoid shorting signals into the internal power and ground plane layers of the substrate 10, anti-pads are typically used in the power plane layers, which compromise power delivery to the memories within the modules.

Double-data rate memory (DDR) uses both the rising and falling edges of a clock signal to operate the memory, resulting in twice the speed of memories using either the leading or the falling edge of the clock signal. DDR3 is the third version of DDR. In DDR3,

as well as other memory types, the command/address bus is a daisy-chained or 'fly-by' bus. This type of bus may have problems with signal integrity due to the unequal routing lengths caused by double-sided memory mounting.

As shown in Figure 1, signal 16a has a first routing length 28a between the connections 20 and 26. The second routing length is between connection 26 on module 12b and ball 30 on module 14a. The second routing length 28b is far shorter than the first. A similar length difference occurs for signal 16b, with the first routing length being from connection 22 to connection 24, and the second routing length being from connection 24 on module 12a to connection 32 on module 14b. It must be noted that connections 20, 22, 24, 26, 30, 32, 34 and 36, are shown here as solder balls, but could be any type of connection used to provide connection between the integrated circuit die and the substrate.

The difference in routing lengths results in non-uniform effective channel impedance. This limits the ability of the devices to receive recognizable signals. Typically, systems are designed to have very regularly and evenly spaced loads and the line is tuned to match the loads. If there is no impedance matching, the signal integrity becomes questionable and higher data rates cannot be supported.

In one approach, the prior art has attempted to perform routing in the semiconductor of the die itself. Connections are redefined to be a different signal using logic in the die to 'mirror' the signal. Logic required in the data path introduces latency in the path, as well as overhead into the device manufacture.

It is possible to interchange the physical connectivity of the various connections at the memory module level, avoiding introduction of logic in the data path, while providing the same benefits as mirroring. An example is shown in Figure 2.

The routing lengths in this embodiment have become uniform. The term 'uniform' as used here does not mean that they are exactly the same, although that may be the case. Uniformity, as used here, means that the impedances in the signal paths are closely matched

to each other between loads. Routing length 1 is now the length from the connection 20 to the connection 30, and routing length 2 is the length from the connection 30 to the connection for the same signal on the next device, not shown. The loads are now more uniformly spaced, which allows the devices to support higher data rates with good signal integrity.

In addition, the configuration shown in Figure 2 has a reduced number of vias. The interchange of signals from the far side of the bottom-side memory modules 12b and 14b have moved to the near side allow the signals to share vias. Reducing the number of vias eases the restrictions on trace spacing, allowing more traces to be placed in a given layer. In turn, this may reduce the number of layers needed in the substrate, reducing costs, as well as mitigating shorts in the power and ground planes.

In order for this approach to be practical, there must be some signals that can be interchanged between the two sides of the module. In order to allow better understanding of the definition of interchangeable signals, it is helpful to discuss how memory modules are typically laid out on a substrate. As shown in Figure 3, a memory controller 38 has a data output of 64 bytes. Each module 40a-40h on the substrate 10 is a X8 module; each one can receive 8 bytes of memory.

It must be noted that in one embodiment of the invention, the interchangeable signals can be selected such that the footprint of the different types of packages can be optimized. As shown in memory module 40a, for example, the data lines have been arranged such that is the package is a X4/X8 package, the data lines are either DQ0-3 for a X4 memory, or DQ0-7 for a X8 package. If the package is a X16 package, all of the data lines are present and DQ0-15 are available for interchangeability. Also, while data lines may be interchangeable within a byte 'lane,' such as DQ[0:7] and DQ[8:15], the adaptability of the interchangeable signals to different package types is enhance if the interchangeability is limited to be within nibbles, such as DQ[0:3], DQ[4:7], DQ[8:11] and DQ[12:15].

Interchangeability actually occurs at the controller 38. The DRAM and the DIMM have no 'knowledge' of what is on any data, whether that data is actually for DQ1 or DQ15. Therefore, these signals are interchangeable. Other types of signals have been identified as being interchangeable, as will be discussed further on. It must be noted that while the data out of the controller is 64 bytes, there are also address and rank-based signals that are sent from the controller in a daisy-chained or 'fly-by' fashion. The signals are passed along a bus and it is the distance between the relevant connections to this bus for which the routing length is desired to be uniform. The interchanging of the signals transmitted from the connections as shown in Figure 2 is what provides this uniformity.

Interchangeability is possible in most memory layouts. For example, the memory system in Figure 4 employs double-sided memory substrates. In addition to the modules 40a-40h on the side facing the viewer, modules such as 42a are on the side away from the viewer. Memory modules arranged in such a system are generally addressed using rank-based signals, such as chip select (CS). In this particular embodiment, the address and rank based signals for the two different ranks are shown coming out of the controller. Generally, rank-based signals are not interchangeable.

Similarly, interchangeability can be performed using 'stacked' memory modules, such as those shown in Figures 5a and 5b. Figure 5a shows an example of a stacked memory arrangement where the connections between the two memory dies 44a and 44b are internal and they use a common connection array such as ball 46. In Figure 5b, each module 44a and 44b have their own external connection array, such as solder balls 46a and 46b. Interchangeability can be employed in this arrangement as well.

Modules are generally arranged as an array of connections, in rows and columns. As will be discussed with regards to Figures 6-8, the arrangement of the connections will be assumed to be in at least four columns, although three columns are possible. This assumption is based upon a typical DRAM layout of 15 or 16 rows of connections by 9 columns.

Generally there are no connections in the middle 3 columns, leaving 3 columns on either 'side' of the module.

In the discussion of interchangeable signals, several different signal abbreviations may be used. These are included with their descriptions in the table below.

Abbreviation	Signal Name	Comments
VSS	Core ground	Usually tied together at substrate
VSSQ	I/O ground	
VDD	Core power	Usually tied together at substrate
VDDQ	I/O power	
RFU	Reserved for future use	
CLK/CLK#	DRAM input clock	
DQ[0:15]	Data signals	Lower and upper bytes (0:7, 8:15)
DQS/DQS#	Strobes for data clock into the DRAM	One for upper and lower bytes
DM	Data mask signal	One for upper and lower bytes
VREFDQ	Voltage reference pin for data	
CS	Chip select	
CKE	Clock enable	
ODT	On-die termination	
RAS	Row address select	
CAS	Column address select	
WE	Write enable	
RST	Reset	
ZQ	Impedance calibration pin	
sCS	Stacked chip select	
sCKE	Stacked clock enable	
sODT	Stacked on-die termination	
sZQ	Stacked impedance calibration pin	
A[0:15]	Address	
BA[0:3]	Bank address	
VREFCA	Voltage reference for command/address	

Interchangeable signals will generally include DQ signals within a byte 'lane' such as DQ[0:7] and DQ[8:15]. Bank addresses, BA[0:3] may be interchangeable. It is possible that BA[2:3] are not present, so only BA[0:1] may be interchanged. Address connections within a row, such as A[3:9] are interchangeable. Generally, VDD and VSS connection locations can be moved around to share vias as well, although this is not shown in the example.

Figure 6 shows a general embodiment of a memory module layout to allow for interchangeable signals. As can be seen, the example is a 16 x 9 array of connections. The array has been arranged in nine columns, with the outer two columns 51, 52, 58 and 59 on either side being identified as for use by interchangeable signals and the middle two columns

53 and 56 being identified as for use by non-interchangeable signals. The outer two columns on each side may be interchangeable providing four interchangeable columns, or only the outer one column on each side providing only two interchangeable columns.

A more specific embodiment of a connection 16 x 9 layout is shown in Figure 7. In this example, VDDQ and VSSQ have had their locations redefined to share vias front to back. Further, A3 and A4, A5 and A6, and A7 and A8 can share vias at the DIMM level. This particular layout has a further advantage that there are only 4 signals per row of the DIMM, allowing better trace separation for the signal traces, further increasing the signal integrity. Reserved for future use (RFU) connections at column 52, rows J and L may be used for sODT, and sCS. Similarly, RFU at column 58, row J could be used for sCKE, and at row D for sZQ.

It is possible to get a 15 x 9 connection layout with some modifications of these rules. If a 2:1 ration of signal to ground pattern can be achieved, and 2 connections are removed, it is possible to achieve a 15 x 9 connection layout with interchangeable connections. An embodiment of this is shown in Figure 8. In this embodiment, the ODT signal has been removed and the bank address 3, BA3 and address 15, A15, share a connection. The RFU connection at column 52, row L becomes BA0.

With these possible layouts in mind, it is helpful to return to Figures 1 and 2. Imagine that the connections 20 and 22 correspond to a location in column 1 of a connection array, and connections 24 and 26 correspond to a location in column 9. In Figure 1, this results in the uneven routing lengths and the extra vias. If the controller were to interchange the signals between columns 1 and 9, moving the signals that had previously been using connection 22 to column 9 and the signals previously using connection 26 to column 1, the result would appear as in Figure 2. The connections 22 and 26 remain in the same place, the signals that had been routed to those columns are interchanged so that connections 20 and 22, and 24 and 26 can be connected together. The availability of interchangeable signals between the outer columns of

the connection array allow for the optimization of the connection layout for double-sided DIMM or other module placement.

Further adjustments and variations on the interchangeable signals are of course possible. For example, the interchangeable signals can also be applied to stacked DRAM technology. In addition, variations of the package type can be employed. For example, the X16 package type may be used, as well as the X4/X8 package type.

Thus, although there has been described to this point a particular embodiment for a method and apparatus for interchangeable connections in a memory module, it is not intended that such specific references be considered as limitations upon the scope of this invention except in-so-far as set forth in the following claims.